

WHITEPAPER

Migrating Unstructured File Data Using Data Dynamics StorageX



Table of Contents

Executive Summary	3
What Is Data Migration?	4
Challenges of NAS Migration	4
Resources	4
Managing Outages	4
Inability to Consolidate or Restructure	4
The Data Dynamics StorageX Solution	5
Architecture	5
Policy-Based Data Movement	6
DFS Integration	6
How StorageX Works	6
Conclusion	7

Executive Summary

A significant portion of enterprise IT budgets is used to keep existing infrastructure running. Increasing operational costs results in a slowdown in innovation and adoption of new technologies. To stay competitive, IT executives are being asked to lower their run rate and promote innovation.

Storage is a major contributor to the growth of the data center footprint. The environmental needs for maintaining a large storage environment drive up power, cooling, and labor costs. Since most IT organizations treat storage as shared infrastructure, managing changes and product lifecycle has become a challenge for storage administrators. The average lifecycle of a storage product is four to five years, which means that every four to five years, storage administrators need to plan on upgrading their estates. This whitepaper explains how to accelerate the adoption and increase the ROI of new NAS hardware using the StorageX platform.

There are various challenges that slow down or prevent the adoption of new NAS technology. Some of these challenges include:

- Getting agreement from business units on cutover outages due to poor prior migration experiences
- Army of resources required to do scripting around traditional host-based data movers
- Lack of automation in planning and design leads to unexpected issues at time of execution
- Limited ability to facilitate restructuring; replication bound by larger containers
- Optimization and restructuring not leveraged as part of the migration; teams are too busy simply trying to move data from source to target
- Inability to move data across platforms without manual or scripted destination creation
- Data movers offer only single protocol support, either CIFS or NFS
- Status and aggregate reporting is basic at best, giving no visibility to business units or technology managers on the progress of the migration
- Abstraction-based solutions used only for migrations are disruptive and difficult to substantiate
- Endless weekends creates “burn out” within implementation teams

To increase the ROI, IT organizations need to look at how to eliminate these challenges and reduce risks and costs of data migration. The NAS footprint is growing at an exponential rate as enterprise organizations become comfortable with the hardware and resiliency provided by the NAS vendors. Storage administrators need to look for opportunities not just to move data, but also to transform data, as they migrate, to decrease burden on data centers.

What Is Data Migration?

Data migration is an expensive and labor-intensive event that occurs every three to five years, depending on the depreciation lifecycle of a storage array. These events, due to their complex nature, consume a significant part of an IT budget. Storage environments are increasingly complex, and this in turn leads to longer and more resource-intensive migration projects. However, these events are not just resource-intensive. They are also disruptive to the users and the business. Managing outages has become a Business As Usual (BAU) function for not just the storage team but also for application and service owners that share the storage infrastructure.

Data migration projects generally take several months and tie up in-house as well as consulting resources. They also put a burden on the application and user communities who have to participate in migration events. It is a common practice for business units to pull out just before a migration event. This leads to further complications for the storage team, as they now have to work around a piece of the infrastructure which is not going to be cut over. Sometimes one business unit pulling out can force the cancellation of the entire migration event, with weeks of planning wasted. This leads to further delay and increased expenses.

NAS migrations are driven by various factors. Most common is the refresh cycle of the hardware. Other factors include introduction of new technology, such as clustered file systems. Datacenter-driven projects and business unit realignments can also drive NAS migrations. These projects have traditionally been done using an army of resources that leverage internal and external replication tools.

Challenges of NAS Migration

We touched briefly on some of the challenges of NAS unstructured data migrations. In this section, we will elaborate further on some of these challenges.

Resources

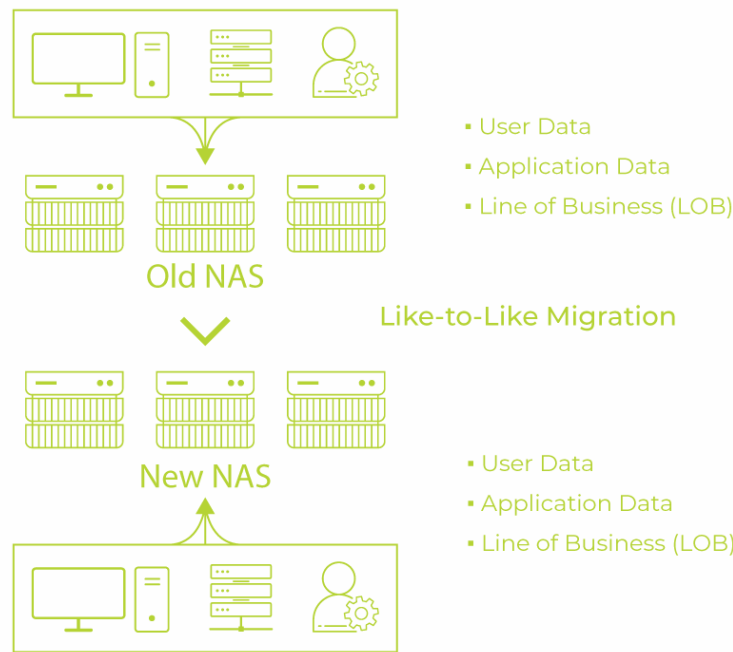
Resources, whether they are people, hardware, or software, are a major challenge. Most projects require various internal resources as well as consulting resources to manage the lifecycle of the migration. Due to the lack of automation, consulting resources are brought on for scripting tasks. Disparate software tools require separate sets of administrative resources who can manage each platform (CIFS or NFS). New hardware platforms are not best utilized, as the goal of a project is to get from point A to B and not to look at how to best use the new platform. Figuring out how to best utilize the platform becomes a post-migration activity.

Managing Outages

One of the biggest challenges in any shared infrastructure migration is coordinating all the applications and business units for the outage. Each application has its own calendar, and coordinating these for an outage is a scheduling nightmare. Performing granular migrations is extremely resource-intensive. They require analysis of the source and detailed mapping of the target, as well as either manual or scripted data movement. Due to the resource-intensive nature of migrations, storage teams generally plan for a migration event pull-out rate of 20%. Sometimes the entire event is cancelled. These delays are what cause 80% of all storage migration projects to go over budget and off schedule.

Inability to Consolidate or Restructure

In the past ten years, the unstructured (NAS) environment in large enterprises has exploded. The NAS environment has become a “dumping ground” for user and application data. Due to challenges with managing the explosive growth rate, storage managers have not been able to segregate data efficiently. User and application data are shared on the same assets. Business realignments and consolidation have taken place, but they are still shared or segregated based on how storage was initially provisioned. Data migrations as we know them now are just a simple move from like to like.



Due to a lack of tools, the main effort is to move from point A to B. A separate effort will take place at some future point in time (if ever) to realign or consolidate. These types of projects drive IT operating expenses up. Why not optimize while you migrate?

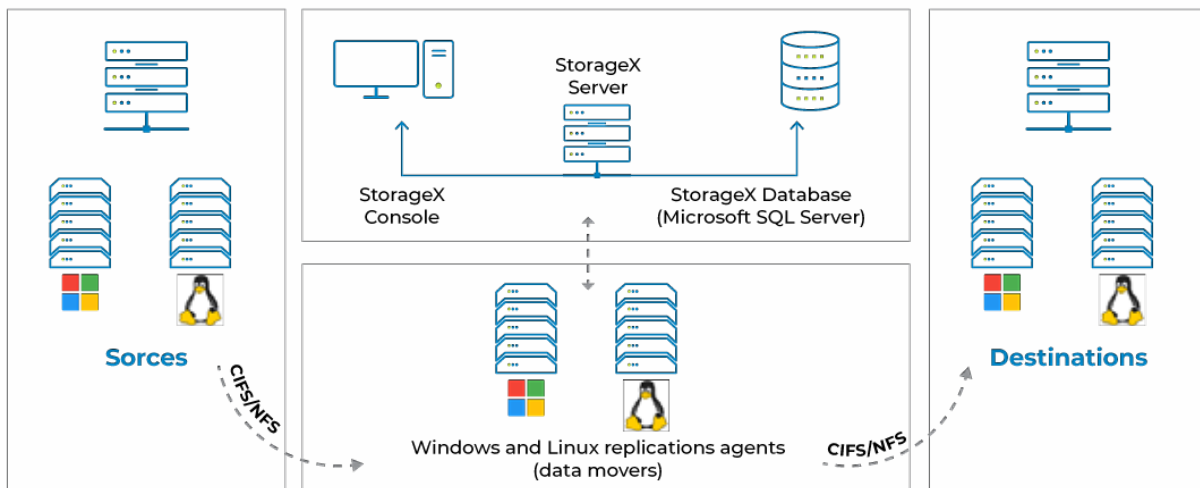
The Data Dynamics StorageX Solution

The Data Dynamics approach to NAS migrations is simple: give storage administrators maximum flexibility when migrating data. Data Dynamics' flagship product, StorageX, gives storage administrators control over their storage environment. Storage administrators can take advantage of the simple user interface and distributed data movement to maximize the migration project.

Storage administrators are no longer bound by legacy containers or to the lines of business for schedules. Storage administrators can also decide on the granularity of the migration. This makes moving to new or heterogeneous storage platforms simple instead of resource-intensive. This also accelerates the adoption of clustered file systems and provides the ability to take advantage of the benefits clustered file systems bring.

Architecture

The StorageX architecture consists of the StorageX server, StorageX Console, StorageX database, which is a Microsoft SQL Server database, and StorageX replication agents, or data movers.



The StorageX server and replication agents can be installed on virtual or physical servers. StorageX replication agents do the heavy lifting, moving file data based on policy and configuration information provided by the StorageX server. The StorageX server and StorageX Console run on Windows 2008 R2 or later. StorageX replication agents run on Windows 2008 R2 or later or Red Hat Enterprise Linux 6 or Red Hat Enterprise Linux 7.

Policy-Based Data Movement

StorageX manages data movement through its policies. Policies can be defined based on various criteria and rules. When used for migration, these policies, combined with distributed replication agents, or data movers, create a powerful data management product. Policies can be configured to migrate data from an entire NAS array. Policies can also be configured in a more granular fashion, at the share or export level. StorageX leverages its available network of replication agents to move the data quickly and efficiently to create a baseline and then continuously copy incremental changes to capture missed (due to files being open) or modified files.

The simple user interface allows the administrator to manage the lifecycle of the replication through one interface, regardless of platform or protocol. Migration policies can be configured to replicate data from any source to any destination, which means exports or shares can be copied to subdirectories for consolidation. Policies can also be configured to move data based on meta-data for archival purposes..

DFS Integration

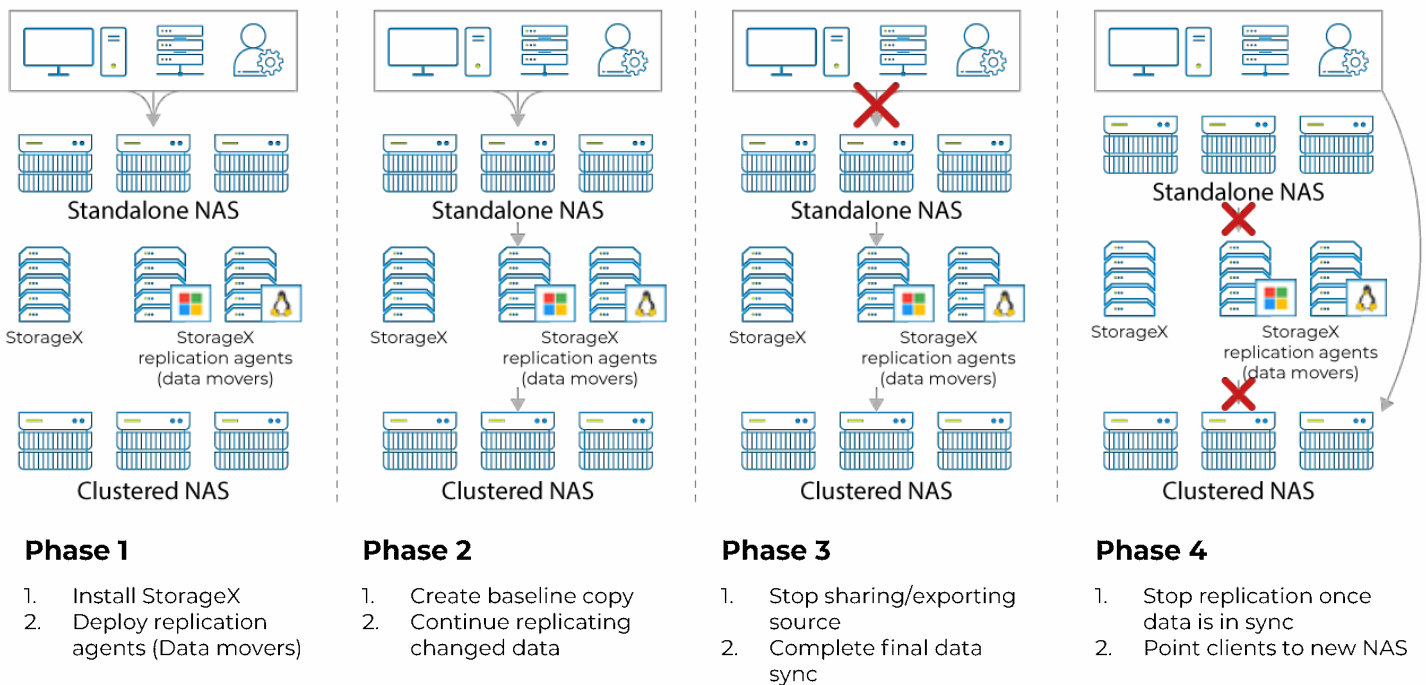
StorageX has its roots as a DFS management product. In DFS environments, StorageX will not only migrate user and group shares, but will also update DFS links to reference the new destination. This creates a powerful tool for seamless cutover of CIFS user and group drives. Storage teams no longer need to engage external resources to script or update DFS links -- it can all be done within one product.

How StorageX Works

We covered earlier the various challenges that face an IT organization when it comes to NAS data migration. In this section, we will show how StorageX helps to address these challenges and mitigate the risk of data migration.

The key components of StorageX all tie together as part of a storage infrastructure. Storage assets and DFS can be viewed and analyzed in one central user interface. A handful of administrators can manage a project without the need for various specialists.

StorageX helps manage the lifecycle of a migration project through this central interface. Resources are defined as sources and potential destinations for migration. Depending on the granularity of the migration, StorageX can analyze the source and recreate the destination based on the source or allow the administrator to define destination creation parameters. Once defined, these parameters will be used for creating the destination. The migration policies then take effect to create a baseline and keep the data on the source as closely in sync with the destination as possible.



Outages are minimized, as administrators can establish replication not just at the array or array container level, but down to the share or export level. This is all done through one user interface, without the need for external tools or scripts. Storage teams don't need to engage platform administrators (Windows/UNIX) in order to copy data at the share or export level. This also facilitates consolidation or restructuring. Users and applications can now be aligned by platform or business.

Storage teams are no longer required to move everything "as is." This also accelerates the adoption of new technologies like clustering. You don't need to move "as is" and then move again to cluster. With StorageX, a non-clustered environment can be entirely moved to a clustered environment. StorageX will recreate all shares and exports as part of the migration. You no longer need an army of resources for any transformational initiatives for unstructured data.

Scheduling is also managed through one interface. If an owner of a share or export for one policy decides not to migrate during a given migration event, that policy can simply be run over and over as needed to maintain incremental copies and then then cut over at a later time. You don't have to move all data at once, and scripts or manual commands do not need to be executed for granular migrations.

Replication agents can be pooled together based on location, platform, or any other criteria. Migration policies will use all replication agents within that pool to maximize the throughput and minimize the outage window for the cutover.

Conclusion

Many of the challenges facing IT organizations today around NAS migrations stem from the lack of tools and automation. Traditionally, IT organizations have worked around this by relying on homogenous, array-based replication tools. However, in today's environment, where flexibility is paramount, those methods do not apply. StorageX gives control back to the IT organization. Storage administrators are empowered by the policy-based StorageX migration engine to not only move data but also to restructure and optimize their file data as part of the migration event. This not only accelerates the adoption of new technology, but also increases the ROI through the use of more granular and flexible migration containers.

Data Dynamics, a global leader in enterprise data management, stands at the forefront of the industry-wide shift towards Digital Trust & Data Democracy. Trusted by 300+ organizations, including 25% of the Fortune 20, the company is recognized for its commitment to creating a transparent, unified, and empowered data ecosystem. Whether addressing data risk, privacy, sovereignty, optimization, sustainability, or facilitating seamless, policy-driven data migration across hybrid and multi-cloud environments, the company is ushering in a new era where data ownership, control, & actionability reside with the data owners.



[Contact Sales](#) [Book a Demo](#)